

USING MACHINE LEARNING TO PREDICT THE WINNER OF A FOOTBALL MATCH

Neeraj Gamare*¹

*¹UG Student, Department Of Information Technology, B. K Birla College Kalyan,
(Empowered Autonomous Status), Kalyan, Maharashtra, India.

ABSTRACT

In this paper, we predict the performance of the Football team will win or lose depends not only on the performance of the team but also on the ground (outside yards), free yards, and last score on this ground, so datasets have been made use for the train model Naïve Bayes, Decision Tree Regression, Support Vector Classifier (SVC), Random Forest Classifier and train all the two datasets. The finding prompted a revolution of the balance between traditional prediction and advanced techniques to find the output of winning team chances. The evaluators the performance of each team player by calculating accuracy and generating confusion matrices classification

Keywords: Artificial Intelligence, Machine Learning.

I. INTRODUCTION

Football, often referred to as "the beautiful game," stands as one of the world's most beloved and passionately followed sports. With a global fan base that transcends boundaries and cultures, football has the unique power to unite people from all walks of life. Whether it's the electrifying roar of the crowd, the skillful dribbles of players, or the heart-stopping goals that grace the pitch, football matches are more than just sporting events; they are an integral part of our collective consciousness. From the colossal stadiums filled with roaring fans to the small neighborhood grounds where dreams take shape, football matches are a showcase of human spirit, competition, and teamwork. These matches are not merely battles for victory but are imbued with stories of underdog triumphs, individual brilliance, and moments that will be etched in history forever. Each football match is a complex interplay of tactical strategies, individual skills, and the unpredictable nature of the game. Teams compete not only for victory but for the hearts and souls of their supporters. Whether it's a local derby or a grand international tournament, the anticipation, excitement, and drama surrounding a football match are unparalleled.

II. METHODOLOGY

In conducting this research, a critical first step was the thorough preprocessing of the dataset. Various data preprocessing steps are performed, truth us handling missing values, dropping accessory columns, creating new features, and checking the dataset must be clean and no null values in datasets. For machine learning some libraries are used like NumPy used for numerical operation, Pandas used for data manipulation and analysis, Matplotlib used for creating plots, Seaborn used for data visualization, and Scikit-learn used for machine learning tasks. Other standard libraries are used for file handling and data processing. Code loads to datasets train matches and train freekicks which contain info about football matches and freekicks. Data loaded into pandas' data frames. The code includes various exploratory data analysis tasks, such as visualizing the number of matches played by each team, analyzing free-kick winners and match winners, and more. For predicting the winner, we evaluate the model like an XGBOOST, Logistics Regression, SVM, and Random Forest.

Classifier. In model evaluation, each model is trained on the training data and evaluated on the test set, after that accuracy score and confusion matrix are calculated to access model performance testing also plays an important role in the model. The data set (presumably for future football matches) is loaded and preprocessed using the same features and engineering steps as the training data.

Machine learning

Machine learning is to predict which team will win a football match based on various features and historical data. The performance of different machine learning models is assessed and compared for this prediction task. Demonstrates this prediction model and provides a basis for selecting the most appropriate model for predicting football match outcomes.

III. RESULTS AND DISCUSSION

Accuracy:

After the model training, many models have a different accuracy of each model.

Table 1. Classification table

MODEL_NAME	ACCURACY
LOGISTICS REGRESSION	0.65
SVM	0.54
RANDOM FOREST	0.64
XG BOOST	0.65

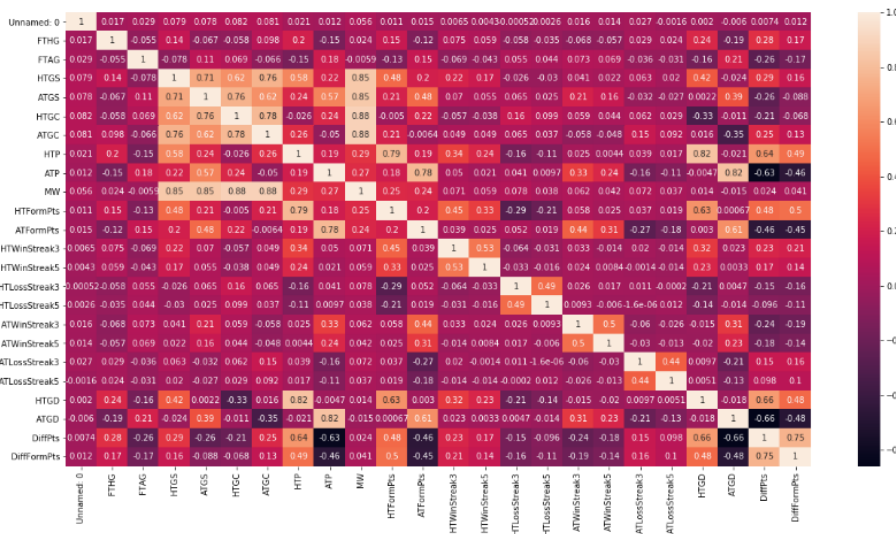


Fig 1: Confusion Matrices

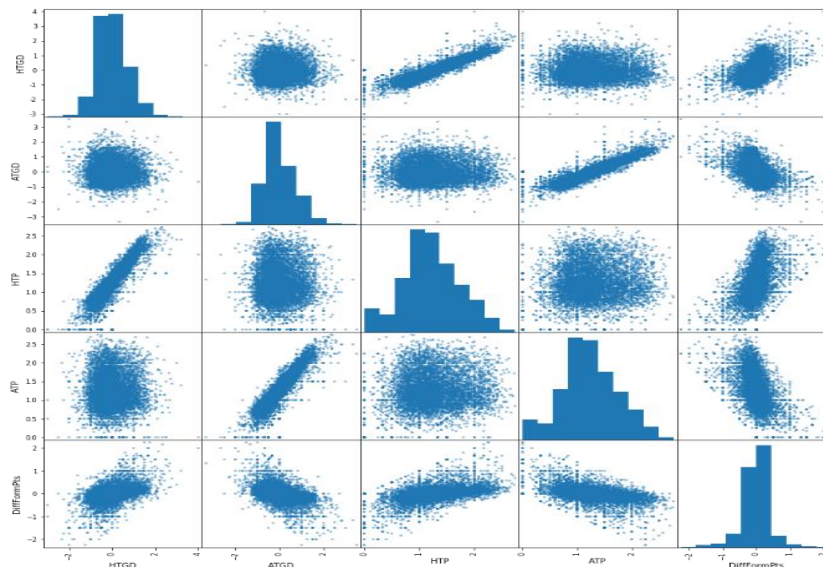


Fig 2: Visualization Of Data Sets

IV. CONCLUSION

In conclusion, it can be inferred that machine learning models have the potential to enhance the accuracy of predicting football match outcomes by considering various influencing factors. The selection of the appropriate model is crucial, as different models may exhibit varying levels of accuracy. This research contributes to the ongoing efforts to combine sports analytics and machine learning to provide valuable insights for football

enthusiasts, analysts, and teams. The paper could benefit from a more detailed discussion of the specific results and implications of the findings. Additionally, references should be provided to support the claims made in the paper. The findings of this research shed light on the potential of machine learning in improving the accuracy of predicting football match outcomes. The models are evaluated in terms of accuracy and confusion matrices, which help visualize the classification results. Notably, each model exhibits different levels of accuracy, emphasizing the significance of selecting the appropriate model for this prediction task.

V. REFERENCES

- [1] R. Beal, S. E. Middleton, T. J. Norman, and S. D. Ramchurn, 'Combining Machine Learning and Human Experts to Predict Match Outcomes in Football: A Baseline Model', Proc. AAAI Conf. Arif. Intell., vol. 35, no. 17, pp. 15447–15451, May 2021, Doi: 10.1609/aaai.v35i17.17815.
- [2] Y. Li, R. Ma, B. Gonçalves, B. Gong, Y. Cui, and Y. Shen, 'Data-driven team ranking and match performance analysis in Chinese Football Super League', Chaos Solitons Fractals, vol. 141, p. 110330, Dec. 2020, Doi: 10.1016/j.chaos.2020.110330.
- [3] Z. Shi, S. Moorthy, and A. Zimmermann, 'Predicting NCAAAB match outcomes using ML techniques – some results and lessons learned'.
- [4] J. Fahey-Gilmour, B. Dawson, P. Peeling, J. Heasman, and B. Rogalski, 'Multifactorial analysis of factors influencing elite Australian football match outcomes: a machine learning approach', Int. J. Compute. Sci. Sport, vol. 18, no. 3, pp. 100–124, Dec. 2019, Doi: 10.2478/ijcss-2019-0020.